

On the Convergence Speed of Spatially Coupled LDPC Ensembles

Vahid Aref^{*†}, Laurent Schmalen[†], and Stephan ten Brink[†],
^{*}School of Information and Communication, EPFL Lausanne, Switzerland
 Email: vahid.aref@epfl.ch
[†]Bell Laboratories, Alcatel-Lucent, Stuttgart, Germany
 Email: {Laurent.Schmalen,Stephan.tenBrink}@alcatel-lucent.com

Abstract—Spatially coupled low-density parity-check codes show an outstanding performance under the low-complexity belief propagation (BP) decoding algorithm. They exhibit a peculiar convergence phenomenon above the BP threshold of the underlying non-coupled ensemble, with a wave-like convergence propagating through the spatial dimension of the graph, allowing to approach the MAP threshold. We focus on this particularly interesting regime in between the BP and MAP thresholds.

On the binary erasure channel, it has been proved [1] that the information propagates with a constant speed toward the successful decoding solution. We derive an upper bound on the propagation speed, only depending on the basic parameters of the spatially coupled code ensemble such as degree distribution and the coupling factor w . We illustrate the convergence speed of different code ensembles by simulation results, and show how optimizing degree profiles helps to speed up the convergence.

Index Terms—Spatially coupled LDPC ensembles, belief propagation, density evolution, convergence speed

I. INTRODUCTION

Low-density parity-check (LDPC) codes are widely used due to their outstanding performance under low-complexity belief propagation (BP) decoding. However, an error probability exceeding that of maximum-a-posteriori (MAP) decoding has to be tolerated with (sub-optimal) BP decoding. However, it has been empirically observed for spatially coupled LDPC (SCLDPC) codes – first introduced by Felström and Zigangirov as convolutional LDPC codes [2] – that the BP performance of these codes can improve dramatically towards the MAP performance of the underlying code under many different settings and conditions, e.g. [3, 4, 5]. This phenomenon – termed *threshold saturation* – has been proven by Kudekar, Richardson and Urbanke in [6, 7]. In particular, they proved that the BP threshold of a coupled LDPC ensemble tends to its MAP threshold on any binary symmetric memoryless channel (BMS). The principle behind threshold saturation seems to be very general and has been applied to a variety of more general scenarios in information theory and computer sciences.

More recently, a new proof technique of threshold saturation has been given by introducing a potential function [8, 9]. Independently, a similar technique has been used in [1] to study one-dimensional continuous coupled systems. It was proven that below the MAP threshold, the information propagates in a wave-like manner with a constant propagation speed by progressing density evolution. This result has also

been extended to discrete spatially coupled one-dimensional systems.

We can distinguish between two convergence regions for SCLDPC codes. If the channel entropy is below the BP threshold of the underlying non-coupled ensemble, the convergence is governed by the ensemble degree distribution. We call this convergence *intra-graph convergence*. If the channel entropy is between the BP and MAP thresholds of the non-coupled ensemble, the wave-like solutions manifests itself after some iterations and the convergence – which we denote *inter-graph convergence* – is dominated by the coupling of the graphs. We particularly focus on this latter convergence phenomenon and are interested in the *inter-graph convergence (propagation) speed* of the wave-like solution. Knowing this speed has several important practical implications: The degree distribution can be optimized in order to maximize the speed for a given target channel and consequently, to minimize the number of required iterations to decode successfully. Additionally, if we employ windowed decoding [10], the convergence speed gives the correct timing to shift the decoding window.

In this paper, we derive upper bounds on the convergence speed of spatially coupled LDPC ensembles for the BEC. Moreover, we compare the convergence speeds of different ensembles and show that the speed is sensitive to the choice of the degree distribution. We also compare the speed of a spatially coupled ensemble on different types of BMS channels.

II. PRELIMINARIES

We briefly explain the graphical structure of LDPC ensembles and spatially coupled LDPC ensemble. Then, we describe the potential functions associated to these ensembles.

A. LDPC Ensembles

LDPC codes are a subset of block codes with sparse parity check matrices. Let n be the block length of the code and m be the number of constraints to satisfy. The design rate of the code is $r = 1 - m/n$. We usually represent LDPC codes by a bipartite graph called factor graph. To each of the n bits, we assign a node, called variable node and we assign a node to each of the m constraints, called check node. We connect variable node i to check node a by an edge if and only if the bit i participates in the corresponding

constraint. To construct a random LDPC code, we sample the degree of node according to a given degree distribution. We represent the degree distribution of the variable nodes by a polynomial $L(x) = \sum_{d=1}^{\infty} L_d x^d$ and the degree distribution of check nodes by $R(x) = \sum_{d=1}^{\infty} R_d x^d$. For a node of degree d , we consider d sockets from which the d edges emanate. Thus, There are $L'(1) = mR'(1)$ sockets in both variable side and check side. After labeling the sockets, we randomly choose a socket from the variable side and connect to a randomly chosen socket in the check side. Finally, we have a random instance of a LDPC(L, R) ensemble. For the detail of construction, we refer to [11].

To study the performance of belief propagation algorithm on sparse graph codes, it is common to use *density evolution*. Let ϵ denote the erasure probability of the channel and let $x^{(t)}$ denote the erasure probability flowing from variable side to the check side at iteration t , then

$$x^{(t+1)} = \epsilon \lambda(1 - \rho(1 - x^{(t)})), \quad (1)$$

where $\lambda(x) = L'(x)/L'(1)$ and $\rho(x) = R'(x)/R'(1)$. We set $x^{(0)} = 1$. The error probability of decoding vanishes if

$$x^{(\infty)} = \lim_{t \rightarrow \infty} x^{(t)} = 0.$$

Definition 1 (BP Threshold). *The BP threshold of the LDPC(L, R) ensemble is defined as:*

$$\epsilon_{\text{BP}} = \sup\{\epsilon \in [0, 1] \mid x^{(\infty)}(\epsilon) = 0\}.$$

The MAP threshold ϵ_{MAP} is defined as the maximum ϵ in which the decoding error probability of MAP decoding is equal to zero. In general, $\epsilon_{\text{BP}} \leq \epsilon_{\text{MAP}}$.

B. Spatially Coupled LDPC Ensemble

We first lay out a set of positions indexed by integers $z \in \mathbb{Z}$ on a line. This line represents a *spatial dimension*. We fix a *coupling factor* which is an integer $w > 1$. Consider $2N$ sets of variable nodes each having n nodes, and locate the sets in positions 1 to $2N$. Similarly, locate $2N + w - 1$ sets of m check nodes each, in positions 1 to $2N + w - 1$. The degree of each variable (check) node is randomly sampled according to the degree distribution $L(x)$ ($R(x)$) leading to $mR'(1)$ sockets at each position.

To construct a random instance of the SCLDPC(L, R, N, w) ensemble, we connect the variable nodes to the check nodes in the following manner: Each of sockets of variable nodes at position $z \in \{1, \dots, 2N\}$ is connected uniformly at random to a socket of check nodes within the range $\{z, \dots, z + w - 1\}$. At the end of this procedure, all $mR'(1)$ sockets of check nodes in position $z \in [w, 2N]$ are occupied except for the check nodes in $z \in [1, w - 1]$, where only a fraction z/w of them are connected. Symmetrically, for the check nodes in $z \in [2N + 1, 2N + w - 1]$, a fraction $(w + 2N - z)/w$ of the sockets are connected. The rest of the sockets are free and we can assume that they are connected to virtual variable nodes

with zero erasure probability. For the detail of construction, we refer to [6].

Let $x_z^{(t)}$ be the erasure probability incoming to check nodes in position $z \in \mathbb{Z}$ at iteration t . The density evolution (DE) equation is

$$x_z^{(t+1)} = \frac{1}{w} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - x_{z+j-k}^{(t)}) \right), \quad (2)$$

where $\epsilon_z = \epsilon$ for $z \in [1, 2N]$, and zero otherwise. We initialize $x_z^{(0)} = 1$ for $z \in [1, 2N + w - 1]$. For the boundary values, $z \notin [1, 2N + w - 1]$, we set $x_z^{(t)} = 0$ for all t .

As density evolution progresses, the perfect boundary information from the left and right sides propagates inward. It was shown in [6, 7] that $x_z^{(t)}$ is non-decreasing sequence for $z \leq N + \lfloor \frac{w-1}{2} \rfloor$ and becomes non-increasing sequence afterward. The symmetric initialization and symmetric boundary conditions induce symmetry on all the erasure probabilities, i.e. $x_z^{(t)} = x_{2N+w-z}^{(t)}$. This system has been termed *two-sided spatially coupled LDPC ensemble*. One half of the spatially coupled ensemble is enough to describe the system.

Definition 2 ([6]). *Let $N' = N + \lfloor \frac{w-1}{2} \rfloor$. The one-sided spatially coupled LDPC ensemble is a modification of (2) defined by fixing $x_z^{(t)} = x_{N'}^{(t)}$ for $z \geq N'$.*

Lemma 1 ([6]). *For the one-sided spatially coupled LDPC ensemble, the densities resulting from density evolution over the BEC satisfy*

$$x_z^{(t)} \geq x_{z-1}^{(t)},$$

and

$$x_z^{(t)} \geq x_z^{(t+1)},$$

for all $z \in \mathbb{Z}$ and $t \geq 0$,

C. Potential Function

We define the potential function of the LDPC(L, R) ensemble as in [8],

$$U(x; \epsilon) = \frac{1}{R'(1)} (1 - R(1 - x)) - x \rho(1 - x) - \frac{\epsilon}{L'(1)} L(1 - \rho(1 - x)) \quad (3)$$

It is indeed equal to the normalized (by $-L'(1)$) replica-symmetric free energies in [12] and it has the following properties.

Lemma 2. *Consider the potential function in (3):*

- i) $\frac{\partial}{\partial x} U(x; \epsilon) = \rho'(1 - x) (x - \epsilon \lambda(1 - \rho(1 - x)))$.
- ii) For $\epsilon < \epsilon_{\text{BP}}$ and $0 < x \leq 1$, $\frac{\partial}{\partial x} U(x; \epsilon) > 0$.
- iii) $U(x; \epsilon)$ is strictly decreasing in terms of ϵ .
- iv) The stationary points of the potential function are the fixed-points of DE equation (1).

Depending on the degree distributions $L(x)$ and $R(x)$, the potential function can have many stationary points. The potential function of regular LDPC(x^3, x^6) ensemble for $\epsilon = 0.475$

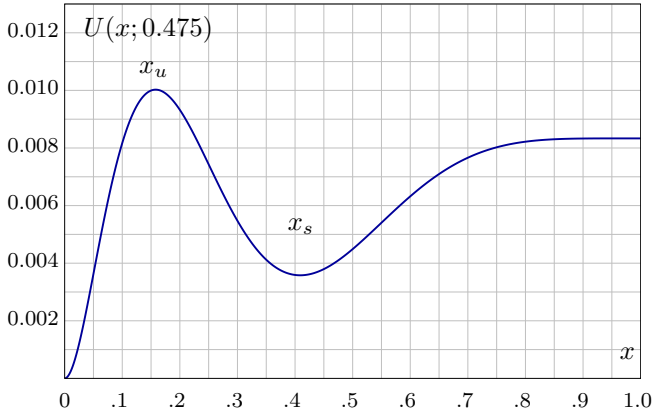


Fig. 1. The potential function of regular LDPC(x^3, x^6) ensemble for $\epsilon = 0.475$. The stable and unstable fixed-points of DE equation are illustrated in the plot.

is depicted in Fig. 1. Each local minimum corresponds to a stable fixed-point of DE equation and each local maximum corresponds to an unstable fixed-point. For $\epsilon < \epsilon_{BP}$ there is only one local minimum which is $x = 0$.

Definition 3. We define the area threshold as

$$\epsilon_{\text{area}} = \sup\{\epsilon \in [0, 1] \mid U(x; \epsilon) \geq 0, \text{ for } x \in [0, 1]\}.$$

It is shown in [8, Lemma 6] that the area threshold is equal to the Maxwell threshold. The Maxwell threshold is equal to MAP threshold on regular ensembles. It has been conjectured that they are equal in general and the equality was recently justified for a large class of LDPC codes in [13]. Thus, $U(x; \epsilon) > 0$ for $\epsilon < \epsilon_{\text{MAP}}$ and becomes zero at ϵ_{MAP} .

Now consider the SCLDPC(L, R, N, w) ensemble. We use the extension of potential function introduced in [8] as follows:

$$U(\mathbf{x}; \epsilon) = \sum_{z=1}^{N'} \frac{1}{R'(1)} (1 - R(1 - x_z)) - x_z \rho(1 - x_z) - \frac{\epsilon}{L'(1)} L \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - x_{z+j}) \right), \quad (4)$$

where $\mathbf{x} = (x_1, \dots, x_{N'})$. We retrieve the DE equation from the partial derivatives of the potential function, i.e.

$$\frac{\partial}{\partial x_z} U(\mathbf{x}; \epsilon) = \rho'(1 - x_z) \times \left(x_z - \frac{1}{w} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - x_{z+j-k}) \right) \right).$$

Thus, the stationary points of $U(\mathbf{x}; \epsilon)$ are the fixed-points of DE equation.

Define the vector \mathbf{y} as $y_z = x_z^{(t)} + h(x_z^{(t+1)} - x_z^{(t)})$ for $0 \leq h \leq 1$. By using the first order Taylor expansion,

$$U(\mathbf{y}; \epsilon) - U(\mathbf{x}^{(t)}; \epsilon) = \sum_{z=1}^{N'} \frac{\partial U(\mathbf{x}^{(t)}; \epsilon)}{\partial x_z} (y_z - x_z^{(t)}) + R_1(\mathbf{y}, \mathbf{x}^{(t)}),$$

where $R_1(\mathbf{y}, \mathbf{x}^{(t)})$ is the remainder term. Define

$$\Delta U_1(\mathbf{y}; \mathbf{x}^{(t)}) = \sum_{z=1}^{N'} \frac{\partial U(\mathbf{x}^{(t)}; \epsilon)}{\partial x_z} (y_z - x_z^{(t)}).$$

Numerical evaluation shows that the remainder term is negative and small in comparison with $\Delta U_1(\mathbf{y}; \mathbf{x}^{(t)})$ and then,

$$U(\mathbf{y}; \epsilon) - U(\mathbf{x}^{(t)}; \epsilon) \approx \Delta U_1(\mathbf{y}; \mathbf{x}^{(t)}).$$

To upper-bound the speed, we must show that there is $\alpha \geq 1$ such that for all t and \mathbf{y} .

$$\alpha \left(U(\mathbf{y}; \epsilon) - U(\mathbf{x}^{(t)}; \epsilon) \right) \leq \Delta U_1(\mathbf{y}; \mathbf{x}^{(t)}). \quad (5)$$

We prove $\alpha \leq 2$ in the next section and in Appendix A for large w . However, our simulation results suggest $\alpha = 1$. The extension to $\alpha = 1$ is currently work in progress.

III. BOUNDS ON THE CONVERGENCE SPEED

Consider the DE equation of one-sided SCLDPC(L, R, N', w) ensemble. Denote the BP threshold and the MAP threshold of the underlying LDPC(L, R) ensemble by ϵ_{BP} and ϵ_{MAP} , respectively. It has been proven in [6, 8] that there is w_0 such that for $\epsilon < \epsilon_{\text{MAP}}$ and $w > w_0$,

$$\lim_{t \rightarrow \infty} \lim_{N' \rightarrow \infty} \lim_{n \rightarrow \infty} x_z^{(t)} = 0,$$

for all $z \in \mathbb{R}$. The question is how fast $x_z^{(t)}$ converges to zero. Assume that $\epsilon_{BP} < \epsilon < \epsilon_{\text{MAP}}$. We distinguish two distinct phases during DE progress. In the first few iterations, all $x_z^{(t)}$ except the ones close to $z = 0$ converges to the forward DE fixed-point of the underlying ensemble, $x_{BP} \neq 0$ (intra-graph convergence). Then, in the next iterations, the information propagates from the boundary $z = 0$ and $x_z^{(t)}$ becomes zero successively (inter-graph convergence).

In this section, we bound the speed of information propagation. First we consider LDPC ensembles whose DE equation has three fixed-points (two stable and one unstable fixed-point). The potential function of one such ensemble is shown in Fig. 1. Many LDPC ensembles including regular LDPC codes have such property. Then we consider the ensembles with more than three DE fixed-points.

A. DE Equation with Three Fixed-points

Fig. 2 shows the solution of DE equation of one-sided SCLDPC($x^3, x^6, 100, 3$) ensemble in different iterations. We observe that the sequence $x_z^{(t)}$, $z \in \mathbb{Z}$ moves uniformly forward with a fixed speed by progressing density evolution. More precisely, there is $T \in \mathbb{N}$ such that

$$x_z^{(t+T)} \leq x_{z-1}^{(t)} < x_z^{(t+T-1)},$$

for all $t > 0$ and $z \in \mathbb{Z}$. The existence of such a DE solution is proven in [1] for the one-sided SCLDPC($L, R, N' = \infty, w$) ensemble on $\epsilon_{BP} < \epsilon < \epsilon_{\text{MAP}}$ in which the DE equation of the underlying ensemble has three fixed-points: zero, an unstable fixed-point x_u and a stable fixed-point $x_s = x_{BP}$, the forward DE fixed-point.

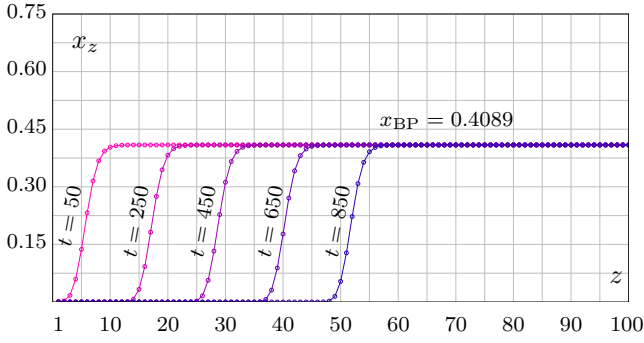


Fig. 2. The solution of DE equation for one-sided SCLDPC($x^3, x^6, 100, 3$) ensemble on BEC with $\epsilon = 0.475$ in different iterations $t = 50, 250, 450, 650$ and 850 . $x_{BP} = 0.4089$. By progressing density evolution, the wave-like solution moves uniformly forward with a fixed speed.

We define the propagation speed of such a DE solution:

Definition 4. For $I \in \mathbb{N}$, define

$$T_I = \min\{T \in \mathbb{N} \mid x_z^{(t+T)} \leq x_{z-I}^{(t)}, \text{ for } t > 0 \text{ and } z \in \mathbb{Z}\}.$$

Additionally, we define the propagation speed $v_I = I/T_I$.

One can show that v_I is an increasing sequence and for

$$\frac{1}{T_1} = v_1 \leq v_I < \frac{1}{T_1 - 1}.$$

We are mostly interested in knowing the speed for ϵ close to ϵ_{MAP} in which the DE solution travels very slowly, i.e. $T_1 \gg 1$. In the following theorems, we upper bound v_1 (or equivalently, lower-bound T_1). We assume that at $t = 0$, the wave-like solution is already formed and for $N' \rightarrow \infty$,

$$\lim_{z \rightarrow N'} x_z^{(t)} = x_{BP},$$

for some iteration t .

Lemma 3. Consider (5). Assume that $T_1 > 1$, then there exists w_0 such that for $w > w_0$, $\alpha \leq 2$.

The sketch of the proof is given in Appendix A.

Theorem 1 (Upper bound). Assume that the DE solution $x_z^{(t)}$ moves forward uniformly and $x_w^{(t)} = 0$ for $t > t_0$. Then, for any $t > t_0$,

$$v_1 \leq B_1 := \frac{\alpha U(x_{BP}; \epsilon)}{\sum_{z \in \mathbb{Z}} \rho'(1 - x_z^{(t)}) (x_z^{(t)} - x_{z-1}^{(t)})^2}, \quad (6)$$

Proof: Appendix B.

As we see later, the above bound is a tight upper bound for large w . For the calculation, we must run DE until the wave-like solution is formed, which in practice, appears after a few iterations. We can additionally state the following (looser) upper bound which eliminates the need to run density evolution:

Theorem 2. Assume that $T_1 > 1$. Let $D = \max_{x \in (0, x_{BP})} \left| \frac{\partial^2}{\partial x^2} U(x; \epsilon) \right|$. Then,

$$B_1 \leq \frac{w \alpha U(x_{BP}; \epsilon)}{2U(x_u; \epsilon) - U(x_{BP}; \epsilon) - D \frac{x_{BP}^2}{w}}. \quad (7)$$

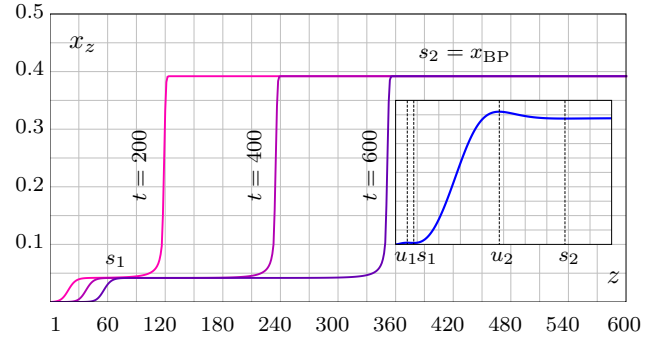


Fig. 3. The DE solution of one-sided SCLDPC($L, R, 600, 3$) ensemble with $L(x) = \frac{153}{283}x^2 + \frac{102}{283}x^3 + \frac{28}{283}x^{51}$ and $R(x) = x^{16}$ on BEC $\epsilon = 0.4$ in different iterations $t = 200, 400, 600$. The inset shows the potential function of LDPC(L, R) for $\epsilon = 0.4$.

which simplifies for $w \rightarrow \infty$ to

$$v_1 \leq B_2 := \frac{w \alpha U(x_{BP}; \epsilon)}{2U(x_u; \epsilon) - U(x_{BP}; \epsilon)}.$$

Proof: Appendix C

Theorem 3 (Lower Bound [1, Theorem 2]). The speed v_1 is lower-bounded by

$$v_1 \geq \frac{w U(x_{BP})}{x_{BP}(1 - \rho(x_{BP}))} - \frac{1}{w}. \quad (8)$$

For $w \gg 1$, $v_1 \geq LB := w U(x_{BP}) / (x_{BP}(1 - \rho(x_{BP})))$.

Note that $U(x_{BP}) / (x_{BP}(1 - \rho(x_{BP})))$ is equal to the area defined in [1].

B. DE Equation with Many DE Fixed-points

In general, the DE equation of irregular LDPC ensembles can have more than three fixed-points in some ϵ . In this case, the DE solution of coupled ensemble can become more complex. It can be a mixture of wave-like solutions with distinct speeds, in which the solution with larger speed overlaps the solutions with lower speed. For instance, consider one-sided SCLDPC($L, R, N' = 600, w = 3$) ensemble with $L(x) = \frac{153}{283}x^2 + \frac{102}{283}x^3 + \frac{28}{283}x^{51}$ and $R(x) = x^{16}$. For $\epsilon_{BP} \approx 0.353 < \epsilon < \epsilon^* \approx 0.399$, the underlying ensemble has three DE fixed-points and thus, the DE solution is as explained in the previous section. For $\epsilon^* < \epsilon < \epsilon_{MAP} \approx 0.403$, the underlying ensemble has five DE fixed-points, namely 0, s_1 and $s_2 = x_{BP}$ the stable fixed-points and u_1 and u_2 the unstable ones. In this case, two wave-like DE solutions appear. In Fig. 3, the DE solution for $\epsilon = 0.4$ is depicted. We observe that for each z , $x_z^{(t)}$ first reduces from s_2 to s_1 by the first wave-like solution and then, reduces from s_1 to 0 by the second wave-like solution.

Let v_1^u and v_1^l denote the propagation speed of the upper and the lower solutions ($v_1^u \geq v_1^l$). v_1^l determines the total number of iterations required for decoding. We can still apply B_1 to upper bound v_1^l with the following assumption. The waves are well-separated such that the right tail of the lower wave converges to s_1 from below and the left tail of upper wave

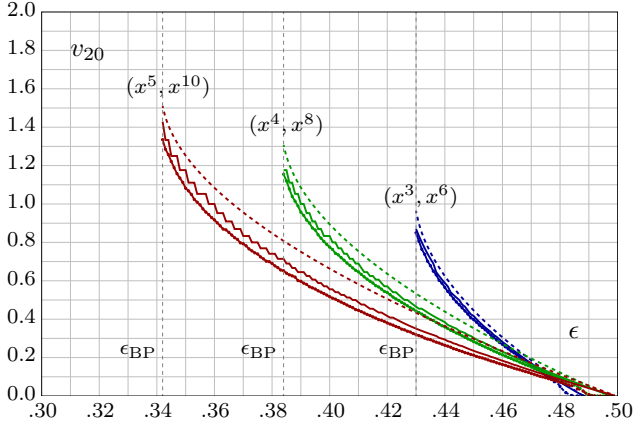


Fig. 4. v_{20} vs. ϵ for some regular SCLDPC($x^k, x^{2k}, N, w = 6$) ensembles and $\epsilon_{BP} < \epsilon < \epsilon_{MAP}$ in each case. v_{20} is computed by DE (solid curves), and also by running BP over code instances with $N = 100$ and $n = 5000$. The results are averaged over 100 instances (dotted curves). The upper bound B_1 is also shown by dashed curves.

converges to s_1 from above (see Fig. 3). If this assumption holds at some t , we can separately study the waves by following the proof of Theorem 1. Shortly,

$$\begin{aligned} \alpha(U(s_2; \epsilon) - U(s_1; \epsilon)) &\geq v_1^u \sum_{x_z \in (s_1, s_2]} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_{z-1}^{(t)})^2 \\ \alpha U(s_1; \epsilon) &\geq v_1^l \sum_{x_z \in (0, s_1]} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_{z-1}^{(t)})^2 \end{aligned}$$

and by summing up both inequalities, and noting that $v_1^u \geq v_1^l$, we have

$$\alpha U(x_{BP}; \epsilon) = \alpha U(s_2; \epsilon) \geq v_1^l \sum_{z \in \mathbb{Z}} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_{z-1}^{(t)})^2.$$

One can also adapt B_2 for the general case by following the similar proof.

IV. SIMULATION RESULTS

As we mentioned before, the numerical results support the conjectured claim $\alpha = 1$. In fact, it is the speed given by the first order Taylor approximation. We therefore set $\alpha = 1$ in the sequel. Consider the SCLDPC(x^3, x^6, N, w) ensemble. We give v_1 and v_{20} for $\epsilon = 0.475$ and different w in Tab. I. We observe that the speed increases almost linearly in terms of w . Note that v_I is an increasing sequence in terms of I and saturates to the real propagation speed. However, for $v_1 \ll 1$, $v_1 \approx v_I$ for any $I \in \mathbb{N}$. In table I, We also compute the upper bounds and the lower bound for each w . We observe that B_1 becomes tighter (to v_I with large I) by increasing w . B_2 seems to be larger than the real speed by an almost constant factor which does not vary significantly with ϵ .

Now consider the asymptotically regular SCLDPC($x^k, x^{2k}, N = 100, w = 6, n = 5000$) ensembles. We plot v_{20} and B_1 for $\epsilon_{BP} < \epsilon < \epsilon_{MAP}$ and for $k = 3, 4$ and 5 in Fig. 4. The curves of B_1 are shown by the dashed lines. We compute v_{20} in two distinct ways: first by the DE equation which are shown by solid lines. We observe that it is a decreasing function and becomes zero at ϵ_{MAP} .

TABLE I
PROPAGATION SPEED OF SCLDPC(x^3, x^6, N, w) ENSEMBLE WITH DIFFERENT w SIZES ON BEC WITH $\epsilon = 0.475$.

w	2	4	8	16	32
v_1	0.035	0.077	0.143	0.250	0.333
v_{20}	0.035	0.079	0.160	0.318	0.625
B_1	0.0503	0.091	0.172	0.333	0.656
B_2	0.435	0.869	1.738	3.477	6.955
LB	0.019	0.038	0.075	0.151	0.302

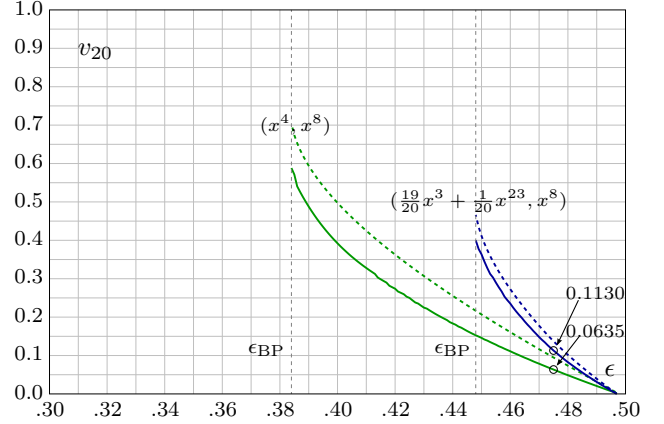


Fig. 5. v_{20} vs. ϵ for regular SCLDPC($x^4, x^8, N, w = 3$) ensemble and for SCLDPC($L(x), x^8, N, w = 3$) ensemble with $L(x) = \frac{19}{20}x^3 + \frac{1}{20}x^{23}$ and $\epsilon_{BP} < \epsilon < \epsilon_{MAP}$ in each case. Both ensembles have the same average variable degree equal to 4. v_{20} is computed by DE (solid curves). The upper bound B_1 is also shown by dashed curves.

Since T_{20} is an integer value and hence, it is not so sensitive to small changes in ϵ , v_{20} is a staircase function. Second, we compute the average v_{20} by running BP over 100 code instances. These curves are shown by dotted lines. As we expect, the empirical v_{20} is close to v_{20} derived by DE. However, the curves deviate from each other for ϵ close to the MAP threshold. The reason is due to the error floor that appears for such ϵ as a result of finite length effects (finite n). The dashed-dotted tail of each curve shows where the error-floor appears.

Figure 4 illustrates that the order of ensembles in terms of speed changes by ϵ . It suggests that for a desired ϵ , we can change the degree distributions in order to obtain a faster convergence speed. To be more fair, we compare two LDPC ensembles with the same average degree. In Fig. 5, we depict v_{20} for the SCLDPC($\frac{19}{20}x^3 + \frac{1}{20}x^{23}, x^8, N, w = 3$) ensemble and for the regular SCLDPC($x^4, x^8, N, w = 3$) ensemble. We observe that the former has much larger speed than the latter over a wide range of ϵ . For instance, the speed is about 1.8 times larger at $\epsilon = 0.475$. However, the regular ensemble has a larger speed for ϵ very close to its ϵ_{MAP} .

These examples show the need for optimizing the speed for a given ϵ . For this purpose, one can use B_1 since it is tight for ϵ close to the MAP threshold.

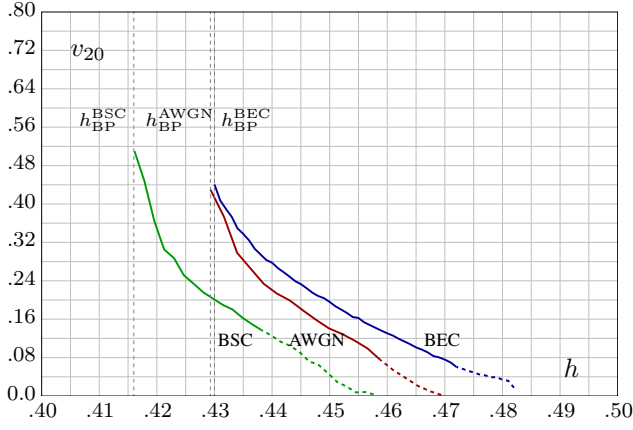


Fig. 6. v_{20} vs. channel entropy h for regular SCLDPC($x^3, x^6, N, w = 3$) ensembles over BEC, BSC and AWGN channel. For each channel, h varies between BP threshold and MAP threshold of underlying ensembles. v_{20} is computed by running BP over code instances with $N = 100$ and $n = 5000$. The results are averaged over 100 instances. The dashed tail of each curve shows where error floor appears. The speed on the BEC is an upper bound over the speed on other channels.

A. Other Types of BMS Channels

Consider the SCLDPC($x^3, x^6, N = 100, 3, n = 5000$) ensemble. Similar to the BEC, simulations results suggest that belief propagation also leads to a wave-like solution on the BSC and on the AWGN channel. In Fig. 6, we compare the empirical speeds for the BEC, BSC and AWGN channel. We compute v_{20} for the channel entropy between BP threshold and MAP threshold in each case. Each point is averaged over 100 code instances. We refer to [11, Chapter 4] for the definition of channel entropy. As discussed before, we observe an error floor close to the MAP threshold due to finite length effects. The dashed tail of each curve indicates where the error-floor appears.

The BP threshold of the underlying ensemble is larger for the BEC than for the AWGN channel and it is furthermore larger for the AWGN channel than for the BSC. The MAP thresholds and the speed preserve the same order. Thus, we conjecture that the upper bound of the speed on BEC is also an upper bound of the speed on the other channels with the same channel entropy.

V. CONCLUSION

In this paper we have analyzed the convergence speed of the wave-like solution that dominates the convergence of spatially coupled codes if the channel entropy lies between the BP and MAP thresholds. Using the Taylor expansion of the spatially coupled code's potential function, we have derived a new upper bound on the convergence speed. This upper bound is based on the degree distribution of the code ensemble and requires the wave-like DE solution. We have additionally derived a looser upper bound that solely depends on the potential function, i.e., on the degree distribution of the ensemble and on the coupling factor. Finally, we have compared the bound with the convergence speeds obtained by density evolution and by simulation of code instances. We have

observed that the upper bound seems to be an upper bound for any binary input memoryless symmetric channel.

APPENDIX A SKETCH OF PROOF OF LEMMA 3

Since $T_1 > 1$,

$$\begin{aligned} x_z^{(t)} - x_z^{(t+1)} &< x_z^{(t)} - x_{z-1}^{(t)} = \frac{1}{w} \left(\epsilon_z \lambda \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1 - x_{z+j}^{(t-1)}) \right) \right. \\ &\quad \left. - \epsilon_{z-w} \lambda \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1 - x_{z+j-w}^{(t-1)}) \right) \right) \\ &\leq \frac{1}{w} \epsilon_z \lambda \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1 - x_{z+j}^{(t-1)}) \right) \leq x_{BP}/w, \end{aligned} \quad (9)$$

where the last inequality is obtained because $x_z^{(t-1)} \leq x_{BP}$ for all z . Now consider y to be $y_z = x_z^{(t)} + h(x_z^{(t+1)} - x_z^{(t)})$ for $0 \leq h \leq 1$. Thus,

$$0 \leq x_z^{(t)} - y_z \leq h x_{BP}/w.$$

We write the second order Taylor expansion of $U(y; \epsilon)$ around $U(x^{(t)}; \epsilon)$:

$$U(y; \epsilon) - U(x^{(t)}; \epsilon) = P(y, x^{(t)}) + R_2(y, x^{(t)}),$$

where $P(\cdot), R_2(\cdot)$ are defined in what follows. We have

$$\begin{aligned} P(y, x^{(t)}) &= \sum_{z=1}^{N'} \frac{\partial U(x^{(t)}; \epsilon)}{\partial x_z} (y_z - x_z^{(t)}) \\ &\quad + \frac{1}{2} \sum_{z=1}^{N'} \sum_{i=1}^{N'} \frac{\partial^2 U(x^{(t)}; \epsilon)}{\partial x_z \partial x_i} (y_z - x_z^{(t)}) (y_i - x_i^{(t)}), \end{aligned}$$

leading to

$$\begin{aligned} P(y, x^{(t)}) &= \sum_{z=1}^{N'} \frac{\partial U(x^{(t)}; \epsilon)}{\partial x_z} (y_z - x_z^{(t)}) \\ &\quad + \frac{1}{2} \sum_{z=1}^{N'} \rho'(1 - x_z^{(t)}) (y_z - x_z^{(t)})^2 - \frac{1}{2} \sum_{z=1}^{N'} Q_z \\ &\quad - \frac{1}{2} \sum_{z=1}^{N'} \rho''(1 - x_z^{(t)}) (x_z^{(t)} - x_z^{(t+1)}) (y_z - x_z^{(t)})^2 \end{aligned}$$

and

$$\begin{aligned} Q_z &= \frac{\rho'(1 - x_z^{(t)}) (y_z - x_z^{(t)})}{w} \sum_{k=0}^{w-1} \epsilon \lambda' \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - x_{z+j-k}^{(t)}) \right) \times \\ &\quad \times \sum_{j=0}^{w-1} \frac{\rho'(1 - x_{z+j-k}^{(t)})}{w} (y_{z+j-k} - x_{z+j-k}^{(t)}). \end{aligned}$$

$R_2(y, x^{(t)})$ is the remainder term. According to Taylor's theorem, there exists $c = x^{(t)} + \gamma(y - x^{(t)})$ for some $0 \leq \gamma \leq 1$ that

$$R_2(y, x^{(t)}) = \frac{1}{6} \sum_{z,i,d} \frac{\partial^3 U(c; \epsilon)}{\partial x_z \partial x_i \partial x_d} (y_z - x_z^{(t)}) (y_i - x_i^{(t)}) (y_d - x_d^{(t)})$$

where $z, i, d \in \{1, \dots, N'\}$. The partial derivatives $\frac{\partial^3}{\partial x_z \partial x_i \partial x_d} U(\mathbf{c}; \epsilon)$, denoted by $U_{zid}(\mathbf{c}; \epsilon)$, are given in equations (11) to (15). By rearranging the terms, we can write $R_2(\mathbf{y}, \mathbf{x}^{(t)})$ in the form of (16) in which there are two positive terms and two negative terms. Now, we show for a large enough w ,

$$U(\mathbf{y}; \epsilon) - U(\mathbf{x}^{(t)}) \leq \frac{1}{2} \sum_{z=1}^{N'} \frac{\partial U(\mathbf{x}^{(t)}; \epsilon)}{\partial x_z} (y_z - x_z^{(t)}). \quad (10)$$

We know that $\frac{\partial}{\partial x_z} U(\mathbf{x}^{(t)}; \epsilon) = \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})$ and then,

$$P(\mathbf{y}, \mathbf{x}^{(t)}) \leq \frac{1}{2} \sum_{z=1}^{N'} \frac{\partial U(\mathbf{x}^{(t)}; \epsilon)}{\partial x_z} (y_z - x_z^{(t)}) - \frac{1}{2} \sum_{z=1}^{N'} Q_z - \frac{1}{2} \sum_{z=1}^{N'} \rho''(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})(y_z - x_z^{(t)})^2$$

We have the result if we show that

$$R_2(\mathbf{y}, \mathbf{x}^{(t)}) - \frac{1}{2} \sum_{z=1}^{N'} Q_z - \frac{1}{2} \sum_{z=1}^{N'} \rho''(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})(y_z - x_z^{(t)})^2 \leq 0. \quad (17)$$

For each z , the second term and the third term of this expression are negative. Furthermore, first term and the last term of $R_2(\mathbf{y}, \mathbf{x}^{(t)})$ in (16) are also negative (note that $y_z - x_z^{(t)} \leq 0$). Finally, we consider the positive terms of $R_2(\mathbf{y}, \mathbf{x}^{(t)})$, which are, for each z ,

$$-\frac{1}{3} \rho''(1 - c_z)(y_z - x_z^{(t)})^3, \quad (18)$$

and

$$-\frac{\rho'(1 - c_z)(y_z - x_z^{(t)})}{6w} \sum_{k=0}^{w-1} \epsilon \lambda'' \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - c_{z+j-k}) \right) \times \left(\sum_{j=0}^{w-1} \frac{\rho(1 - c_{z+j-k})}{w} (y_{z+j-k} - x_{z+j-k}^{(t)}) \right)^2. \quad (19)$$

Now compare (18) and

$$\frac{1}{2} \rho''(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})(y_z - x_z^{(t)})^2. \quad (20)$$

For a large w the above term dominates (18) since

$$x_z^{(t)} - h x_{\text{BP}}/w \leq y_z \leq c_z \leq x_z^{(t)}. \quad (21)$$

Now consider the pair of Q_z and (19) for each z . We have the bound

$$\left| \sum_{j=0}^{w-1} \frac{\rho'(1 - c_{z+j-k})}{w} (y_{z+j-k} - x_{z+j-k}^{(t)}) \right| < h \frac{x_{\text{BP}}}{w} \sum_{j=0}^{w-1} \frac{\rho'(1 - c_{z+j-k})}{w} < h \frac{x_{\text{BP}}}{w} \rho'(1).$$

Q_z is positive and $O(\frac{1}{w^2})$ for a large w but (19) is $O(\frac{1}{w^3})$ and thus, there exists w' such that for $w > w'$, Q_z becomes larger than (19).

Consolidating the above observations provides the desired inequality (17) for a large w and hence, (10) follows.

APPENDIX B PROOF OF THEOREM 1

Recall that $x_z^{(t)}$ is an increasing sequence in terms of z and a decreasing sequence in terms of t . We assume that the DE solution moves uniformly, i.e. there is $T \in \mathbb{N}$ such that

$$x_z^{(t+T)} \leq x_{z-1}^{(t)} < x_z^{(t+T-1)},$$

for all $t \in \mathbb{N}$ and $z \in \mathbb{Z}$. Without loss of generality, assume that $t = 0$ and define $y_z = x_z^{(0)}$. Thus,

$$y_z = x_z^{(0)} \geq x_z^{(1)} \geq \dots \geq x_z^{(T-1)} > x_{z-1}^{(0)} = y_{z-1} \geq x_z^{(T)}. \quad (22)$$

Let $\mathbf{x}^{(t)}$ denote the sequence $x_z^{(t)}$ at iteration t and let $\mathbf{y}^{[k]}$ denote the k -shifted sequence $y_z^{[k]} = y_{z-k}$ for $k \in \mathbb{Z}$. According to Lemma 3,

$$\alpha(U(\mathbf{x}^{(t+1)}; \epsilon) - U(\mathbf{x}^{(t)}; \epsilon)) \leq - \sum_{z=1}^{N'} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})^2.$$

Thus, we can write

$$U(\mathbf{y}^{[1]}; \epsilon) - U(\mathbf{y}^{[0]}; \epsilon) = \sum_{t=0}^{T-2} U(\mathbf{x}^{(t+1)}; \epsilon) - U(\mathbf{x}^{(t)}; \epsilon) + U(\mathbf{y}^{[1]}; \epsilon) - U(\mathbf{x}^{(T-1)}; \epsilon),$$

and hence

$$\alpha(U(\mathbf{y}^{[1]}; \epsilon) - U(\mathbf{y}^{[0]}; \epsilon)) \leq - \sum_{t=0}^{T-2} \sum_{z=1}^{N'} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})^2 - \sum_{z=1}^{N'} \rho'(1 - x_z^{(T-1)})(x_z^{(T-1)} - y_{z-1})(x_z^{(T-1)} - x_z^{(T)}),$$

where in the last application of Lemma 3, we have used $h = (x_z^{(T-1)} - y_{z-1})/(x_z^{(T-1)} - x_z^{(T)})$ since we assume that the DE solution moves uniformly. Since $y_{z-1} \geq x_z^{(T)}$, we have

$$\alpha(U(\mathbf{y}^{[1]}; \epsilon) - U(\mathbf{y}^{[0]}; \epsilon)) \leq - \sum_{t=0}^{T-2} \sum_{z=1}^{N'} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})^2 - \sum_{z=1}^{N'} \rho'(1 - x_z^{(T-1)})(x_z^{(T-1)} - y_{z-1})^2,$$

or equivalently,

$$\alpha(U(\mathbf{y}^{[0]}; \epsilon) - U(\mathbf{y}^{[1]}; \epsilon)) \geq \sum_{t=0}^{T-2} \sum_{z=1}^{N'} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})^2 + \sum_{z=1}^{N'} \rho'(1 - x_z^{(T-1)})(x_z^{(T-1)} - y_{z-1})^2,$$

Now consider the right hand side of the above inequality and let us exchange the order of the sums. We can change it even

$$U_{zzz}(\mathbf{c}; \epsilon) = \rho'''(1-c_z) \left(c_z - \frac{1}{w} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1-c_{z+j-k}) \right) \right) - 2\rho''(1-c_z) + 3 \frac{\rho''(1-c_z)\rho'(1-c_z)}{w^2} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda' \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1-c_{z+j-k}) \right) - \frac{(\rho'(1-c_z))^3}{w^3} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda'' \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1-c_{z+j-k}) \right), \quad (11)$$

$$U_{zzi}(\mathbf{c}; \epsilon) = \frac{\rho''(1-c_z)\rho'(1-c_i)}{w^2} \sum_{k=\max(0, z-i)}^{w-1+\min(0, z-i)} \epsilon_{z-k} \lambda' \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1-c_{z+j-k}) \right) - \frac{(\rho'(1-c_z))^2 \rho'(1-c_i)}{w^3} \sum_{k=\max(0, z-i)}^{w-1+\min(0, z-i)} \epsilon_{z-k} \lambda'' \left(\frac{1}{w} \sum_{j=0}^{w-1} 1 - \rho(1-c_{z+j-k}) \right), \quad (12)$$

$$U_{zzi}(\mathbf{c}; \epsilon) = U_{ziz}(\mathbf{c}; \epsilon) = U_{izz}(\mathbf{c}; \epsilon), \quad (13)$$

$$U_{zid}(\mathbf{c}; \epsilon) = - \frac{\rho'(1-c_z)\rho'(1-c_i)\rho'(1-c_d)}{w^3} \sum_{k=\max(0, z-i, z-d)}^{w-1+\min(0, z-i, z-d)} \epsilon_{z-k} \lambda'' \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1-c_{z+j-k}) \right), \quad (14)$$

$$U_{zid}(\mathbf{c}; \epsilon) = U_{zdi}(\mathbf{c}; \epsilon) = U_{dzi}(\mathbf{c}; \epsilon) = U_{diz}(\mathbf{c}; \epsilon) = U_{izz}(\mathbf{c}; \epsilon) = U_{idz}(\mathbf{c}; \epsilon) \quad (15)$$

$$R_2(\mathbf{y}, \mathbf{x}^{(t)}) = \frac{1}{6} \sum_{z=1}^{N'} \rho'''(1-c_z) \left(c_z - \frac{1}{w} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1-c_{z+j-k}) \right) \right) (y_z - x_z^{(t)})^3 - \frac{1}{3} \sum_{z=1}^{N'} \rho''(1-c_z) (y_z - x_z^{(t)})^3 - \frac{1}{6} \sum_{z=1}^{N'} \frac{\rho'(1-c_z)(y_z - x_z^{(t)})}{w} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda'' \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1-c_{z+j-k}) \right) \times \left(\sum_{j=0}^{w-1} \frac{\rho'(1-c_{z+j-k})}{w} (y_{z+j-k} - x_{z+j-k}^{(t)}) \right)^2 + \frac{3}{6} \sum_{z=1}^{N'} \frac{\rho''(1-c_z)(y_z - x_z^{(t)})^2}{w} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda' \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1-c_{z+j-k}) \right) \times \left(\sum_{j=0}^{w-1} \frac{\rho'(1-c_{z+j-k})}{w} (y_{z+j-k} - x_{z+j-k}^{(t)}) \right). \quad (16)$$

for $N' \rightarrow \infty$ because the summands are positive. Then for each z , we minimize the following summation: The application of Jensen's inequality leads to

$$\sum_{t=0}^{T-2} \rho'(1-x_z^{(t)})(x_z^{(t)} - x_z^{(t+1)})^2 + \rho'(1-x_z^{(T-1)})(x_z^{(T-1)} - y_{z-1})^2 = \sum_{t=0}^{T-1} \rho'(1-x_z^{(t)}) l_t^2 \quad \frac{1}{T} \sum_{t=0}^{T-1} l_t^2 \geq \left(\frac{1}{T} \sum_{t=0}^{T-1} l_t \right)^2 = \frac{1}{T^2} (y_z - y_{z-1})^2,$$

where $l_t = x_z^{(t)} - x_z^{(t+1)}$ for $0 \leq t < T-1$ and $l_{T-1} = x_z^{(T-1)} - y_{z-1}$. Since $\rho'(1-x)$ is a decreasing function, we have

$$\sum_{t=0}^{T-1} \rho'(1-x_z^{(t)}) l_t^2 \geq \rho'(1-y_z) \sum_{t=0}^{T-1} l_t^2.$$

Due to (22), $l_t \geq 0$ and

$$\sum_{t=0}^{T-1} l_t = x_z^{(0)} - y_{z-1} = y_z - y_{z-1}.$$

and thus,

$$\alpha(U(\mathbf{y}^{[0]}; \epsilon) - U(\mathbf{y}^{[1]}; \epsilon)) \geq \frac{1}{T} \sum_{z=1}^{N'} \rho'(1-y_z) (y_z - y_{z-1})^2. \quad (23)$$

Considering the definition of the potential function in (4), we observe that $U(\mathbf{y}^{[0]}; \epsilon) - U(\mathbf{y}^{[1]}; \epsilon)$ is a telescoping sum with only the difference between the first and the last term

remaining, i.e.

$$\begin{aligned}
U(\mathbf{y}^{[0]}; \epsilon) - U(\mathbf{y}^{[1]}; \epsilon) &= \frac{1}{R'(1)} (1 - R(1 - y_{N'})) - \\
&- y_{N'} \rho(1 - y_{N'}) - \frac{1}{R'(1)} (1 - R(1 - y_0)) + y_0 \rho(1 - y_0) - \\
&- \frac{\epsilon}{L'(1)} L \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - y_{N'+j}) \right) - \\
&+ \frac{\epsilon}{L'(1)} L \left(1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - y_j) \right).
\end{aligned}$$

With $y_z = 0$ for $z \leq w$ (by the theorem assumption) and $y_z = y_{N'} = x_{\text{BP}}$ for $z > N'$, we have

$$U(\mathbf{y}^{[0]}; \epsilon) - U(\mathbf{y}^{[1]}; \epsilon) = U(x_{\text{BP}}; \epsilon). \quad (24)$$

Finally, the statement of the theorem is concluded from (23) and (24).

APPENDIX C PROOF OF THEOREM 2

In order to prove the theorem, we require the following useful lemma.

Lemma 4. *Let $x_z^{(t)}$ be the solution of DE equation (2) at some iteration t . Further assume that $T_1 > 1$. Then,*

$$x_z^{(t)} - x_{z-1}^{(t)} \geq \left| \frac{x_z^{(t)} - \epsilon \lambda(1 - \rho(1 - x_z^{(t)}))}{w} \right|, \quad (25)$$

for $z \leq N'$.

Proof: First we prove that $x_z^{(t)} - x_{z-1}^{(t)} \geq (x_z^{(t)} - \epsilon \lambda(1 - \rho(1 - x_z^{(t)}))) / w$. Define

$$\nu_z^{(t)} := 1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - x_{z+j}^{(t)}).$$

Note that $\nu_z^{(t)}$ is an increasing sequence in terms of z , because $1 - \rho(1 - x)$ is an increasing function and $x_z^{(t)}$ is an increasing sequence in z . Thus, according to (2),

$$x_z^{(t)} = \frac{1}{w} \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda(\nu_{z-k}^{(t-1)}) \leq \epsilon_z \lambda(\nu_z^{(t-1)}),$$

and by rewriting the above equation,

$$\begin{aligned}
(w-1)x_z^{(t)} &\geq \sum_{k=0}^{w-1} \epsilon_{z-k} \lambda(\nu_{z-k}^{(t-1)}) - \epsilon_z \lambda(\nu_z^{(t-1)}) \\
&= \sum_{k=1}^{w-1} \epsilon_{z-k} \lambda(\nu_{z-k}^{(t-1)}) = \sum_{k=0}^{w-2} \epsilon_{z-1-k} \lambda(\nu_{z-1-k}^{(t-1)})
\end{aligned}$$

Thus, we can upper bound $w x_{z-1}^{(t)}$ as

$$\begin{aligned}
w x_{z-1}^{(t)} &= \sum_{k=0}^{w-1} \epsilon_{z-1-k} \lambda(\nu_{z-1-k}^{(t-1)}) \\
&= \sum_{k=0}^{w-2} \epsilon_{z-1-k} \lambda(\nu_{z-1-k}^{(t-1)}) + \epsilon_{z-w} \lambda(\nu_{z-w}^{(t-1)}) \\
&\leq (w-1)x_z^{(t)} + \epsilon_{z-w} \lambda(\nu_{z-w}^{(t-1)}) \\
&\leq (w-1)x_z^{(t)} + \epsilon \lambda(1 - \rho(1 - x_{z-1}^{(t-1)})) \\
&\leq (w-1)x_z^{(t)} + \epsilon \lambda(1 - \rho(1 - x_z^{(t)})). \quad (26)
\end{aligned}$$

The last two inequalities result from the facts that $\epsilon_{z-w} \leq \epsilon$ and that

$$\nu_{z-w}^{(t-1)} = 1 - \frac{1}{w} \sum_{j=0}^{w-1} \rho(1 - x_{z-w+j}^{(t-1)}) \leq 1 - \rho(1 - x_{z-1}^{(t-1)}).$$

Furthermore, $x_{z-1}^{(t-1)} < x_z^{(t)}$ since $T_1 > 1$. The result is given by rearrangement of (26). Similarly, we prove that $x_z^{(t)} - x_{z-1}^{(t)} \geq (\epsilon \lambda(1 - \rho(1 - x_z^{(t)})) - x_z^{(t)}) / w$.

$$x_{z-1}^{(t)} = \frac{1}{w} \sum_{k=0}^{w-1} \epsilon_{z-1-k} \lambda(\nu_{z-1-k}^{(t-1)}) \geq \epsilon_{z-w} \lambda(\nu_{z-w}^{(t-1)}),$$

and thus,

$$\begin{aligned}
(w-1)x_{z-1}^{(t)} &\leq \sum_{k=0}^{w-1} \epsilon_{z-1-k} \lambda(\nu_{z-1-k}^{(t-1)}) - \epsilon_{z-w} \lambda(\nu_{z-w}^{(t-1)}) \\
&= \sum_{k=0}^{w-2} \epsilon_{z-1-k} \lambda(\nu_{z-1-k}^{(t-1)}) = \sum_{k=1}^{w-1} \epsilon_{z-k} \lambda(\nu_{z-k}^{(t-1)}).
\end{aligned}$$

For $x_z^{(t)}$, we have

$$\begin{aligned}
w x_z^{(t)} &= \sum_{k=1}^{w-1} \epsilon_{z-k} \lambda(\nu_{z-k}^{(t-1)}) + \epsilon_z \lambda(\nu_z^{(t-1)}) \\
&\geq (w-1)x_{z-1}^{(t)} + \epsilon_z \lambda(\nu_z^{(t-1)}) \\
&\geq (w-1)x_{z-1}^{(t)} + \epsilon \lambda(1 - \rho(1 - x_z^{(t-1)})) \\
&\geq (w-1)x_{z-1}^{(t)} + \epsilon \lambda(1 - \rho(1 - x_z^{(t)})).
\end{aligned}$$

We conclude the result since

$$(w+1)x_z^{(t)} \geq w x_{z-1}^{(t)} + \epsilon \lambda(1 - \rho(1 - x_z^{(t)})).$$

■

We now turn to the proof of Theorem 2. To derive an upper bound on the convergence speed, we must lower-bound the following summation:

$$\sum_{z=1}^{N'} \rho'(1 - x_z^{(t)}) (x_z^{(t)} - x_{z-1}^{(t)})^2. \quad (27)$$

Let $y_z = x_z^{(t)}$ for $z \in \mathbb{Z}$. From Lemma 2 and Lemma 4, we have

$$\rho'(1 - y_z)(y_z - y_{z-1}) \geq \left| \frac{\partial U(y_z; \epsilon)}{w \partial y_z} \right| =: \left| \frac{U'(y_z; \epsilon)}{w} \right|.$$

We can divide the interval $[0, x_s]$ into four sub-intervals (see also Fig. 1):

- i) $x \in [0, x_1]$ where $U'(x; \epsilon) \geq 0$ and $U''(x; \epsilon) \geq 0$.
- ii) $x \in (x_1, x_u]$ where $U'(x; \epsilon) \geq 0$ and $U''(x; \epsilon) \leq 0$.
- iii) $x \in (x_u, x_2]$ where $U'(x; \epsilon) \leq 0$ and $U''(x; \epsilon) \leq 0$.
- iv) $x \in (x_2, x_s]$ where $U'(x; \epsilon) \leq 0$ and $U''(x; \epsilon) \geq 0$.

Let a denote the index where $y_a \leq x_1 < y_{a+1}$. Similarly, we define b and c for x_u and x_2 , respectively. We split (27) into four sum according to the sub-intervals and lower-bound them separately.

i) For $z \leq a$, since $U''(y_z; \epsilon) \geq 0$ (\cup -convexity), $U'(y_z; \epsilon)(y_z - y_{z-1}) \geq U(y_z; \epsilon) - U(y_{z-1}; \epsilon)$. Thus,

$$\sum_{z=1}^a |U'(y_z; \epsilon)|(y_z - y_{z-1}) \geq U(y_a; \epsilon).$$

ii) Since the potential functions becomes \cap -convex for $a+1 \leq z \leq b$,

$$U(y_z; \epsilon) - U(y_{z-1}; \epsilon) \leq U'(y_{z-1}; \epsilon)(y_z - y_{z-1}).$$

According to mean value theorem, $\exists k_z \in [y_{z-1}, y_z]$ such that

$$U'(y_z; \epsilon) = U'(y_{z-1}; \epsilon) + U''(k_z; \epsilon)(y_z - y_{z-1}).$$

Let $D = \max_{x \in [0, x_s]} |U''(x; \epsilon)|$. Then,

$$\begin{aligned} U'(y_z; \epsilon)(y_z - y_{z-1}) &\geq U'(y_{z-1}; \epsilon)(y_z - y_{z-1}) - D(y_z - y_{z-1})^2 \\ &\geq U(y_z; \epsilon) - U(y_{z-1}; \epsilon) \\ &\quad - D \frac{x_s}{w} (y_z - y_{z-1}). \end{aligned}$$

The last inequality is due to $y_z - y_{z-1} \leq x_s/w$ which can be shown by writing the DE equation for y_z and y_{z-1} and then, since $y_z - y_{z-1}$ is a telescoping sum, the inequality follows. Hence,

$$\begin{aligned} \sum_{z=a+1}^b |U'(y_z; \epsilon)|(y_z - y_{z-1}) &\geq U(y_b; \epsilon) - U(y_a; \epsilon) \\ &\quad - D \frac{x_s}{w} (y_b - y_a) \end{aligned}$$

iii) Similar to i) we now use the \cap -convexity property. For $b+1 < z \leq c$, $U'(y_z; \epsilon)(y_z - y_{z-1}) \leq U(y_z; \epsilon) - U(y_{z-1}; \epsilon)$ since $U''(y_z; \epsilon) \leq 0$ and then,

$$\begin{aligned} \sum_{z=b+1}^c |U'(y_z; \epsilon)|(y_z - y_{z-1}) &= - \sum_{z=b+1}^c U'(y_z; \epsilon)(y_z - y_{z-1}) \\ &\geq U(y_b; \epsilon) - U(y_c; \epsilon). \end{aligned}$$

iv) For $z > c$, $U''(y_z; \epsilon) \geq 0$, which implies

$$U'(y_{z-1}; \epsilon)(y_z - y_{z-1}) \leq U(y_z; \epsilon) - U(y_{z-1}; \epsilon).$$

By the mean value theorem, we get

$$U'(y_z; \epsilon) \leq U'(y_{z-1}; \epsilon) + D(y_z - y_{z-1}),$$

and consequently,

$$U'(y_z; \epsilon)(y_z - y_{z-1}) \leq U(y_z; \epsilon) - U(y_{z-1}; \epsilon) + D \frac{x_s}{w} (y_z - y_{z-1}),$$

and,

$$\begin{aligned} \sum_{z=c+1}^{N'} |U'(y_z; \epsilon)|(y_z - y_{z-1}) &= - \sum_{z=c+1}^{N'} U'(y_z; \epsilon)(y_z - y_{z-1}) \\ &\geq U(y_c; \epsilon) - U(y_{N'}; \epsilon) \\ &\quad - D \frac{x_s}{w} (y_{N'} - y_c), \end{aligned}$$

where $y_{N'} = x_s$. By summing up all terms,

$$\begin{aligned} \sum_{z=1}^{N'} |U'(y_z; \epsilon)|(y_z - y_{z-1}) &\geq 2U(y_b; \epsilon) - U(x_s; \epsilon) \\ &\quad - D \frac{x_s}{w} (x_s - y_c + y_b - y_a). \end{aligned}$$

by definition $y_b \leq x_u < y_{b+1}$. Since $U'(x_u; \epsilon) = 0$, there exists $k \in [y_b, x_u]$ such that,

$$\begin{aligned} U(y_b; \epsilon) &= U(x_u; \epsilon) + \frac{U''(k; \epsilon)}{2} (x_u - y_b)^2 \\ &\geq U(x_u; \epsilon) - \frac{D}{2} (x_u - y_b)^2 \\ &\geq U(x_u; \epsilon) - \frac{D}{2} \frac{x_s}{w} (x_u - y_b). \end{aligned}$$

The last inequality holds since $x_u - y_b \leq y_{b+1} - y_b \leq x_s/w$. Finally,

$$\begin{aligned} \sum_{z=1}^{N'} |U'(y_z; \epsilon)|(y_z - y_{z-1}) &\geq 2U(x_u; \epsilon) - U(x_s; \epsilon) \\ &\quad - D \frac{x_s}{w} (x_s - y_c + x_u - y_a) \\ &\geq 2U(x_u; \epsilon) - U(x_s; \epsilon) - D \frac{x_s^2}{w}. \end{aligned}$$

We have the result by noting that

$$\sum_{z=1}^{N'} \rho'(1 - x_z^{(t)})(x_z^{(t)} - x_{z-1}^{(t)})^2 \geq \frac{1}{w} \sum_{z=1}^{N'} |U'(y_z; \epsilon)|(y_z - y_{z-1}).$$

ACKNOWLEDGMENT

We thank Rüdiger Urbanke and Nicolas Macris for insightful discussions. Vahid Aref was supported by grant No. 200021-125347 of the Swiss National Science Foundation. Laurent Schmalen was supported by the German Government in the frame of the CELTIC+/BMBF project SASER-SaveNet.

REFERENCES

- [1] S. Kudekar, T. Richardson, and R. Urbanke, "Wave-like solutions of general one-dimensional spatially coupled systems," *preprint, arXiv:1208.5273*, 2012.
- [2] A. J. Felström and K. S. Zigangirov, "Time-varying periodic convolutional codes with low density parity check matrix," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 2181–2190, 1999.
- [3] M. Lentmaier, A. Sridharan, K. S. Zigangirov, and D. J. Costello, "Terminated LDPC convolutional codes with thresholds close to capacity," in *Proc. ISIT*, 2005, pp. 1372–1376.
- [4] M. Lentmaier, A. Sridharan, D. J. Costello, Jr., and K. S. Zigangirov, "Iterative decoding threshold analysis for LDPC convolutional codes," *IEEE Trans. Inform. Theory*, vol. 56, no. 10, 2010.
- [5] M. Lentmaier, D. G. M. Mitchell, G. P. Fettweis, and D. J. Costello, "Asymptotically regular LDPC codes with linear distance growth and

thresholds close to capacity,” in *Proc. Inform. Theory and Appl. Workshop (ITA)*, January 2010, pp. 1–8.

- [6] S. Kudekar, T. J. Richardson, and R. L. Urbanke, “Threshold saturation via spatial coupling: why convolutional LDPC ensembles perform so well over the BEC,” *IEEE Trans. Inform. Theory*, vol. 57, no. 2, 2009.
- [7] S. Kudekar, T. Richardson, and R. Urbanke, “Spatially coupled ensembles universally achieve capacity under belief propagation,” *preprint, arXiv:1201.2999*, Jan. 2012.
- [8] A. Yedla, Y.-Y. Jian, P. S. Nguyen, and H. D. Pfister, “A simple proof of threshold saturation for coupled scalar recursions,” in *Proc. Intl. Symp. on Turbo Codes and Iter. Inform. Proc. (ISTC)*, 2012, pp. 51–55.
- [9] S. Kumar, A. J. Young, N. Macris, and H. D. Pfister, “A proof of threshold saturation for spatially-coupled LDPC codes on BMS channels,” *preprint, arXiv:1301.6111*, 2013.
- [10] A. Iyengar, P. Siegel, R. Urbanke, and J. Wolf, “Windowed decoding of spatially coupled codes,” *IEEE Trans. Inform. Theory*, vol. 59, no. 4, pp. 2277–2292, Apr. 2013.
- [11] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.
- [12] N. Macris, “Griffith–Kelly–Sherman correlation inequalities: A useful tool in the theory of error correcting codes,” *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 664–683, 2007.
- [13] A. Giurgiu, N. Macris, and R. L. Urbanke, “Spatial coupling as a proof technique,” *preprint, arXiv:1301.5676*, 2013.